RAW SEQUENCE LISTING

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) no errors detected.

Application Serial Number: 10/033,129 B

Source: TFW16

Date Processed by STIC: 1-3-05

ENTERED

RAW SEQUENCE LISTING   DATE: 01/03/2005
PATENT APPLICATION: US/10/033,129B   TIME: 14:56:25

Input Set : **A:\8325-2001-30seqlistcomplete.txt**
Output Set: **N:\CRF4\01032005\J033129B.raw**

```
 3 <110> APPLICANT: CHOO, Yen
 4        KLUG, Aaron
 5        SANCHEZ-GARCIA, Isidro
 7 <120> TITLE OF INVENTION: BINDING PROTEINS FOR RECOGNITION OF DNA
 9 <130> FILE REFERENCE: 8325-2001.30 (G1-US3)
11 <140> CURRENT APPLICATION NUMBER: 10/033,129B
12 <141> CURRENT FILING DATE: 2001-12-27
14 <150> PRIOR APPLICATION NUMBER: 08/793,408
15 <151> PRIOR FILING DATE: 1997-06-02
17 <150> PRIOR APPLICATION NUMBER: PCT/GB95/01949
18 <151> PRIOR FILING DATE: 1995-08-17
20 <160> NUMBER OF SEQ ID NOS: 125
22 <170> SOFTWARE: PatentIn version 3.2
24 <210> SEQ ID NO: 1
25 <211> LENGTH: 60
26 <212> TYPE: DNA
27 <213> ORGANISM: Artificial
29 <220> FEATURE:
30 <223> OTHER INFORMATION: oligonucleotide
32 <400> SEQUENCE: 1
33 ctcctgcagt tggacctgtg ccatggccgg ctgggccgca tagaatggaa caactaaagc      60
36 <210> SEQ ID NO: 2
37 <211> LENGTH: 92
38 <212> TYPE: PRT
39 <213> ORGANISM: Artificial
41 <220> FEATURE:
42 <223> OTHER INFORMATION: zinc finger protein used in phage display library
45 <220> FEATURE:
46 <221> NAME/KEY: MISC_FEATURE
47 <222> LOCATION: (48)..(51)
48 <223> OTHER INFORMATION: Xaa = any amino acid
50 <220> FEATURE:
51 <221> NAME/KEY: MISC_FEATURE
52 <222> LOCATION: (53)..(54)
53 <223> OTHER INFORMATION: Xaa = any amino acid
55 <220> FEATURE:
56 <221> NAME/KEY: MISC_FEATURE
57 <222> LOCATION: (56)..(56)
58 <223> OTHER INFORMATION: Xaa = any amino acid
60 <220> FEATURE:
61 <221> NAME/KEY: MISC_FEATURE
62 <222> LOCATION: (57)..(57)
63 <223> OTHER INFORMATION: Xaa = Arg or Lys
```

RAW SEQUENCE LISTING   DATE: 01/03/2005
PATENT APPLICATION: US/10/033,129B   TIME: 14:56:25

Input Set : A:\8325-2001-30seqlistcomplete.txt
Output Set: N:\CRF4\01032005\J033129B.raw

```
 65 <400> SEQUENCE: 2
 67 Met Ala Glu Glu Arg Pro Tyr Ala Cys Pro Val Glu Ser Cys Asp Arg
 68 1               5                  10                  15
 71 Arg Phe Ser Arg Ser Asp Glu Leu Thr Arg His Thr Arg Ile His Thr
 72              20                  25                  30
```
W-->
```
 75 Gly Gln Lys Pro Phe Gln Cys Arg Ile Cys Met Arg Asn Phe Ser Xaa
 76          35                  40                  45
 79 Xaa Xaa Xaa Leu Xaa Xaa His Xaa Xaa Thr His Thr Gly Glu Lys Pro
 80      50                  55                  60
 83 Phe Ala Cys Asp Ile Cys Gly Arg Lys Phe Ala Arg Ser Asp Glu Arg
 84 65                  70                  75                  80
 87 Lys Arg His Thr Lys Ile His Leu Arg Gln Lys Asp
 88              85                  90
 91 <210> SEQ ID NO: 3
 92 <211> LENGTH: 26
 93 <212> TYPE: DNA
 94 <213> ORGANISM: Artificial
 96 <220> FEATURE:
 97 <223> OTHER INFORMATION: oligonucleotide
 99 <400> SEQUENCE: 3
100 tatgacttgg atgggagacc gcctgg                                    26
103 <210> SEQ ID NO: 4
104 <211> LENGTH: 28
105 <212> TYPE: DNA
106 <213> ORGANISM: Artificial
108 <220> FEATURE:
109 <223> OTHER INFORMATION: oligonucleotide
111 <400> SEQUENCE: 4
112 aattccaggc ggtctcccat ccaagtca                                  28
115 <210> SEQ ID NO: 5
116 <211> LENGTH: 21
117 <212> TYPE: DNA
118 <213> ORGANISM: Artificial
120 <220> FEATURE:
121 <223> OTHER INFORMATION: oligonucleotide
123 <400> SEQUENCE: 5
124 tatatagcgt gggcgtatat a                                         21
127 <210> SEQ ID NO: 6
128 <211> LENGTH: 24
129 <212> TYPE: DNA
130 <213> ORGANISM: Artificial
132 <220> FEATURE:
133 <223> OTHER INFORMATION: oligonucleotide
135 <400> SEQUENCE: 6
136 gcgtatatac gcccacgcta tata                                      24
139 <210> SEQ ID NO: 7
140 <211> LENGTH: 21
141 <212> TYPE: DNA
142 <213> ORGANISM: Artificial
```

```
144 <220> FEATURE:
145 <223> OTHER INFORMATION: oligonucleotide
148 <220> FEATURE:
149 <221> NAME/KEY: misc_feature
150 <222> LOCATION: (10)..(12)
151 <223> OTHER INFORMATION: n is a, c, g, or t
153 <400> SEQUENCE: 7
```
W--> 154 tatatagcgn nngcgtatat a   21
```
157 <210> SEQ ID NO: 8
158 <211> LENGTH: 24
159 <212> TYPE: DNA
160 <213> ORGANISM: Artificial
162 <220> FEATURE:
163 <223> OTHER INFORMATION: oligonucleotide
166 <220> FEATURE:
167 <221> NAME/KEY: misc_feature
168 <222> LOCATION: (13)..(15)
169 <223> OTHER INFORMATION: n is a, c, g, or t
171 <400> SEQUENCE: 8
```
W--> 172 gcgtatatac gcnnncgcta tata   24
```
175 <210> SEQ ID NO: 9
176 <211> LENGTH: 33
177 <212> TYPE: DNA
178 <213> ORGANISM: Artificial
180 <220> FEATURE:
181 <223> OTHER INFORMATION: BCR-ABL
183 <400> SEQUENCE: 9
184 ttccatggag acgcagaagc ccttcagcgg cca                    33
187 <210> SEQ ID NO: 10
188 <211> LENGTH: 33
189 <212> TYPE: DNA
190 <213> ORGANISM: Artificial
192 <220> FEATURE:
193 <223> OTHER INFORMATION: BCR
195 <400> SEQUENCE: 10
196 ttccatggag acgcaggtga gttcctcacg cca                    33
199 <210> SEQ ID NO: 11
200 <211> LENGTH: 33
201 <212> TYPE: DNA
202 <213> ORGANISM: Artificial
204 <220> FEATURE:
205 <223> OTHER INFORMATION: ABL
207 <400> SEQUENCE: 11
208 cccctttctc ttccagaagc ccttcagcgg cca                    33
211 <210> SEQ ID NO: 12
212 <211> LENGTH: 33
213 <212> TYPE: PRT
214 <213> ORGANISM: Artificial
216 <220> FEATURE:
```

RAW SEQUENCE LISTING   DATE: 01/03/2005
PATENT APPLICATION: US/10/033,129B   TIME: 14:56:25

Input Set : A:\8325-2001-30seqlistcomplete.txt
Output Set: N:\CRF4\01032005\J033129B.raw

```
217 <223> OTHER INFORMATION: 1A
219 <400> SEQUENCE: 12
221 Met Ala Glu Glu Lys Pro Phe Gln Cys Arg Ile Cys Met Arg Asn Phe
222 1               5                   10                  15
225 Ser Asp Arg Ser Ser Leu Thr Arg His Thr Arg His Thr Gly Glu Lys
226             20                  25                  30
229 Pro
233 <210> SEQ ID NO: 13
234 <211> LENGTH: 33
235 <212> TYPE: PRT
236 <213> ORGANISM: Artificial
238 <220> FEATURE:
239 <223> OTHER INFORMATION: 1B
241 <400> SEQUENCE: 13
243 Met Ala Glu Glu Lys Pro Phe Gln Cys Arg Ile Cys Met Arg Asn Phe
244 1               5                   10                  15
247 Ser Glu Arg Gly Thr Leu Ala Arg His Glu Lys His Thr Gly Glu Lys
248             20                  25                  30
251 Pro
255 <210> SEQ ID NO: 14
256 <211> LENGTH: 27
257 <212> TYPE: PRT
258 <213> ORGANISM: Artificial
260 <220> FEATURE:
261 <223> OTHER INFORMATION: 2A
263 <400> SEQUENCE: 14
265 Phe Gln Cys Arg Ile Cys Met Arg Asn Phe Ser Gln Gly Gly Asn Leu
266 1               5                   10                  15
269 Val Arg His Leu Arg His Thr Gly Glu Lys Pro
270             20                  25
273 <210> SEQ ID NO: 15
274 <211> LENGTH: 26
275 <212> TYPE: PRT
276 <213> ORGANISM: Artificial
278 <220> FEATURE:
279 <223> OTHER INFORMATION: 3A
281 <400> SEQUENCE: 15
283 Phe Gln Cys Arg Ile Cys Met Arg Asn Phe Ser Gln Ala Gln Thr Leu
284 1               5                   10                  15
287 Gln Arg His Leu Lys His Thr Gly Glu Lys
288             20                  25
291 <210> SEQ ID NO: 16
292 <211> LENGTH: 26
293 <212> TYPE: PRT
294 <213> ORGANISM: Artificial
296 <220> FEATURE:
297 <223> OTHER INFORMATION: 3B
299 <400> SEQUENCE: 16
301 Phe Gln Cys Arg Ile Cys Met Arg Asn Phe Ser Gln Ala Ala Thr Leu
```

```
302 1                 5                      10                    15
305 Gln Arg His Leu Lys His Thr Gly Glu Lys
306               20                    25
309 <210> SEQ ID NO: 17
310 <211> LENGTH: 26
311 <212> TYPE: PRT
312 <213> ORGANISM: Artificial
314 <220> FEATURE:
315 <223> OTHER INFORMATION: 3C
317 <400> SEQUENCE: 17
319 Phe Gln Cys Arg Ile Cys Met Arg Asn Phe Ser Gln Ala Gln Asp Leu
320 1                 5                      10                    15
323 Gln Arg His Leu Lys His Thr Gly Glu Lys
324               20                    25
327 <210> SEQ ID NO: 18
328 <211> LENGTH: 89
329 <212> TYPE: PRT
330 <213> ORGANISM: Artificial
332 <220> FEATURE:
333 <223> OTHER INFORMATION: zinc finger polypeptide
335 <400> SEQUENCE: 18
337 Met Ala Glu Glu Lys Pro Phe Gln Cys Arg Ile Cys Met Arg Asn Phe
338 1                 5                      10                    15
341 Ser Asp Arg Ser Ser Leu Thr Arg His Thr Arg Thr His Thr Gly Glu
342               20                    25                    30
345 Lys Pro Phe Gln Cys Arg Ile Cys Met Arg Asn Phe Ser Asp Arg Ser
346          35                    40                    45
349 His Leu Thr Arg His Thr Arg Thr His Thr Gly Glu Lys Pro Phe Gln
350       50                    55                    60
353 Cys Arg Ile Cys Met Arg Asn Phe Ser Asp Arg Ser Asn Leu Thr Arg
354 65                    70                    75                    80
357 His Thr Arg Thr His Thr Gly Glu Lys
358                    85
361 <210> SEQ ID NO: 19
362 <211> LENGTH: 10
363 <212> TYPE: PRT
364 <213> ORGANISM: Artificial
366 <220> FEATURE:
367 <223> OTHER INFORMATION: zinc finger binding motif
369 <400> SEQUENCE: 19
371 Arg Ser Asp His Leu Thr Thr His Ile Arg
372 1                 5                      10
374 <210> SEQ ID NO: 20
375 <211> LENGTH: 10
376 <212> TYPE: PRT
377 <213> ORGANISM: Artificial
379 <220> FEATURE:
380 <223> OTHER INFORMATION: zinc finger binding motif
382 <400> SEQUENCE: 20
```

RAW SEQUENCE LISTING ERROR SUMMARY   DATE: 01/03/2005
PATENT APPLICATION: US/10/033,129B   TIME: 14:56:26

Input Set : A:\8325-2001-30seqlistcomplete.txt
Output Set: N:\CRF4\01032005\J033129B.raw

Please Note:
Use of n and/or Xaa have been detected in the Sequence Listing. Please review the
Sequence Listing to ensure that a corresponding explanation is presented in the <220>
to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:2; Xaa Pos. 48,49,50,51,53,54,56,57
Seq#:7; N Pos. 10,11,12
Seq#:8; N Pos. 13,14,15

Invalid <213> Response:
Use of "Artificial" only as "<213> Organism" response is incomplete,
per 1.823(b) of New Sequence Rules. Valid response is Artificial Sequence.

Seq#:1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20,21,22,23,24,25,26,27
Seq#:28,29,30,31,32,33,34,35,36,37,38,39,40,41,42,43,44,45,46,47,48,49,50,51
Seq#:52,53,54,55,56,57,58,59,60,61,62,63,64,65,66,67,68,69,70,71,72,73,74,75
Seq#:76,77,78,79,80,81,82,83,84,85,86,87,88,89,90,91,92,93,94,95,96,97,98,99
Seq#:100,101,102,103,104,105,106,107,108,109,110,111,112,113,114,115,116,117
Seq#:118,119,120,121,122,123,124,125

VERIFICATION SUMMARY   DATE: 01/03/2005
PATENT APPLICATION: **US/10/033,129B**   TIME: 14:56:26

Input Set : **A:\8325-2001-30seqlistcomplete.txt**
Output Set: **N:\CRF4\01032005\J033129B.raw**

L:75 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:2 after pos.:32
M:341 Repeated in SeqNo=2
L:154 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:7 after pos.:0
L:172 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8 after pos.:0